



# A Data-Driven Fuzzy Approach for Predicting the Remaining Useful Life in Dynamic Failure Scenarios of a Nuclear Power Plant

F. Di Maio, Enrico Zio

## ► To cite this version:

F. Di Maio, Enrico Zio. A Data-Driven Fuzzy Approach for Predicting the Remaining Useful Life in Dynamic Failure Scenarios of a Nuclear Power Plant. Reliability Engineering and System Safety, 2010, 95 (1), pp.49-57. 10.1016/j.ress.2009.08.001 . hal-00610470

**HAL Id: hal-00610470**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-00610470>**

Submitted on 16 Jan 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A DATA-DRIVEN FUZZY APPROACH FOR PREDICTING THE REMAINING USEFUL LIFE IN DYNAMIC FAILURE SCENARIOS OF A NUCLEAR SYSTEM

Enrico Zio, Francesco Di Maio

*Energy Department, Polytechnic of Milan*

*Via Ponzio 34/3, 20133 Milano, Italy*

[enrico.zio@polimi.it](mailto:enrico.zio@polimi.it)

## ABSTRACT

*This paper presents a similarity-based approach for prognostics of the Remaining Useful Life (RUL) of a system, i.e. the lifetime remaining between the present and the instance when the system can no longer perform its function. Data from failure dynamic scenarios of the system are used to create a library of reference trajectory patterns to failure. Given a failure scenario developing in the system, the remaining time before failure is predicted by comparing by fuzzy similarity analysis its evolution data to the reference trajectory patterns and aggregating their times to failure in a weighted sum which accounts for their similarity to the developing pattern. The prediction on the failure time is dynamically updated as time goes by and measurements of signals representative of the system state are collected. The approach allows for the on-line estimation of the RUL.*

*For illustration, a case study is considered regarding the estimation of RUL in failure scenarios of the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS).*

**Key Words:** Prognostics, Residual Useful Life, Recovery Time, Emergency Accident Management, Operator Support System, Nuclear Power Plant, Lead-Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS), Pointwise Fuzzy Similarity Analysis.

## 1. Introduction

Existing Nuclear Power Plants (NPPs) strive to improve safety, maintain availability and reduce operation and maintenance costs. Moreover, power uprates and life extensions increase the needs of techniques for diagnosing and prognosing the NPPs health, because of the more probable occurrence of component degradation and failure as load is increased, or changed, and age advances.

In generality, prognostics is an important and challenging task in Reliability, Availability, Maintainability and Safety (RAMS). The primary goal of a prognostic system is to indicate whether the structure, system or component (SSC) of interest can perform its function throughout its lifetime with reasonable assurance and, in case it cannot, to estimate the Remaining Useful Life (RUL), i.e. the lifetime remaining before it can no longer perform its function [Jardine et al., 2006].

Approaches to prognostics can be categorized broadly into model-based and data-driven [Chiang et al., 2001]. Model-based prognostics attempts to incorporate physical models of the system into the estimation of the RUL. However, uncertainty due to the assumptions and simplifications of the adopted models may pose significant limitations on this approach. On the contrary, data-driven techniques utilize monitored operational data related to system health. They can be beneficial when understanding of first principles of system operation is not straightforward or when the system is so complex that developing an accurate model is prohibitively expensive. Data-driven approaches can often be deployed quickly and cheaply, and still provide wide coverage of system behavior. An added value of data-driven techniques is their ability to transform high-dimensional noisy data into lower dimensional information useful for decision-making [Dragomir et al., 2007].

In practice, when the cost/benefit ratio of using physics-based degradation evolution models is not favorable but sufficient data (possibly simulated) are available for constructing a map of the damage space, data-driven techniques may be more apt to the prognostics task. Furthermore, recent

advances in sensor technology and refined simulation capabilities enable us to continuously monitor the health of operating components and manage the related large amount of reference data.

Data-driven approaches can be divided into two categories: statistical techniques (regression methods, ARMA models, etc.) and Artificial Intelligence (AI) techniques (neural networks, fuzzy systems, etc.).

The most natural data-driven technique for RUL estimation is to fit a curve of the available data of the component degradation evolution using regression models and extrapolate the curve to the criteria indicating failure. However, the available component history sometimes may be short and incomplete, so that extrapolation may lead to large errors. The same problem arises when employing ARMA models, although the method can handle also situation in which more run-to-failure data are unavailable or insufficient [Yan et al., 2004].

With respect to AI techniques, the most commonly used prediction methods are based on Neural Networks [Peel et al., 2008; Barlett et al., 1992; Santosh et al., 2009]. For prognostic tasks, the most promising methods are Recurrent Neural Networks (RNNs) [Campolucci et al., 1999] and Neuro-Fuzzy (NF) systems [Wang et al., 2004]. In spite of the recognized power of neural network modeling techniques, limitations still exist which prevent their wide use in safety critical applications, e.g. in nuclear technology. Skepticism against their use in such application relates to the lack of a systematic approach for selecting the structure and parameters of the model and to their black-box character which limits intuition with respect to the understanding of their performance [Wang et al., 2008].

An opportunity for increased transparency and openness of data-driven model is offered by fuzzy logic methods, which are increasingly proposed in modern diagnostic technologies. Based on the principles of Zadeh's fuzzy set theory, fuzzy logic provides a formal mathematical framework for dealing with the vagueness of everyday reasoning [Zadeh, 1965]. As opposed to binary

reasoning based on ordinary set theory, within the fuzzy logic framework measurement uncertainty and estimation imprecision can be properly accommodated [Yuan et al., 1997; Zio et al., 2005].

The goal of this work is to develop a prognostics tool to be embedded in an operator support system for emergency accident management. In this context, the remaining useful life represents the recovery time available to the operator for deciding on the corrective action to avoid system failure. In emergency situations, poor decisions may be taken because of the short time available for sorting out the relevant information and lack of expert knowledge [Glasstone et al., 1998]; on the contrary, timely and correct decisions can prevent an incident from developing into a severe accident or mitigate its undesired consequences.

The paper presents the computational framework devised for integrating on-line monitored data for system RUL prediction. A set of failure trajectory patterns (hereafter called reference trajectory patterns) is collected in a reference library and a fuzzy based data-driven similarity analysis is performed for predicting the remaining life of a newly developing failure trajectory (hereafter called test trajectory pattern). A novelty of the proposal lies in the reliance on a fuzzy definition of trajectory pattern similarity to capture and integrate the ambiguous information carried by the measured signal. More specifically, the pattern matching process is based on a fuzzy distance evaluation algorithm between the reference and test trajectory patterns [Angstenberger, 2001].

An application is presented with reference to dynamic failure scenarios of the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) with digital Instrumentation and Control (I&C) [Cammi et al., 2006].

The paper is structured as follows. Section 2 provides a detailed description of the algorithm for the fuzzy data similarity comparison underpinning the prognostic approach to RUL estimation. Section 3 presents the mechanistic model of the LBE-XADS. In Section 4, the results of the

application of the approach to LBE-XADS accident scenarios are presented. Finally, some conclusions are drawn in Section 5.

## 2. Methodology

It is assumed that  $N$  trajectories of evolution in time of values of relevant signals (reference trajectory patterns) are available from measurements collected in dynamic failure scenarios of the system of interest. These trajectories last all the way to system failure, i.e., to the time when anyone of the signals reaches the threshold value beyond which the system loses its functionality.

A failure scenario is developing in the system, whose signals trajectory in time (test trajectory pattern) is monitored.

The idea underpinning the RUL estimation is to evaluate the similarity between the test trajectory pattern of the developing failure scenario and the  $N$  reference trajectory patterns stored in the database and use the RULs of these latter to estimate the RUL of the former, accounting for how similar they are [Angstenberger, 2001].

Figure 1 shows a schematic sketch of the computational framework; for ease of illustration, a single signal  $f(t)$  is considered (a generalization to the case of multidimensional trajectories is straightforward).

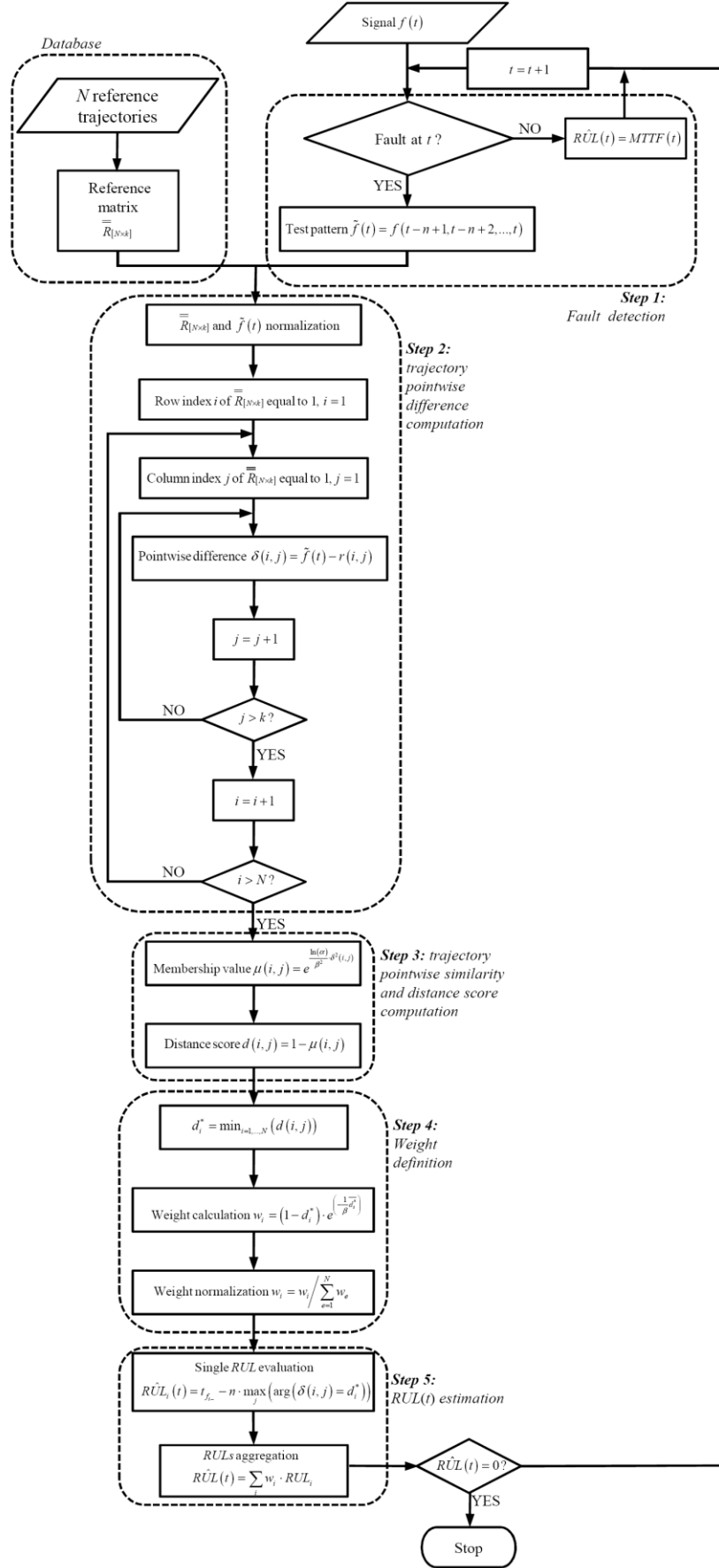


Figure 1 The flowchart of the fuzzy based data-driven approach

In what follows, the steps of the procedure are reported in details:

- **Step 1: fault detection.** The signal  $f(t)$  is continuously monitored throughout the time horizon of observation  $T$ , starting from (discrete) time  $t=1$ ; at each time  $t$ , its value is recorded and appended to the vector of the values collected at the time steps preceding  $t$ . For reasons which will become clear in the following, the database constituting the reference trajectory pattern library is organized in a reference matrix  $\overline{\mathbf{R}}_{[N \times k]}$ , where  $k = \frac{T}{n}$ , whose generic element  $r(i, j)$  is the  $j$ -th segment of length  $n$  of the values of the  $i$ -th reference trajectory,  $i = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, k$  (Figures 2-3), normalized in the range  $[0.2, 0.8]$ . As long as no signal deviation is detected, the system is qualified as working in nominal conditions and the estimate  $\hat{RUL}(t)$  of the remaining useful life made at the generic time  $t$  is taken equal to the system Mean Time to Failure  $MTTF(t)$ , obtained from the remaining useful life  $RUL_i(t)$  of all the failure trajectories in the reference library:

$$\hat{RUL}(t) = MTTF(t) = \frac{1}{(|i| |t_{f_i} > t|)} \sum_{i|t_{f_i} > t} (t_{f_i} - t) = \frac{1}{(|i| |t_{f_i} > t|)} \sum_{i|t_{f_i} > t} RUL_i(t) \quad (1)$$

where  $t_{f_i}$  is the system failure time along the  $i$ -th trajectory (i.e., the time when the signal value exceeds the threshold beyond which the system loses its functionality),  $(|i| |t_{f_i} > t|)$  is the cardinality of the set of reference trajectories whose failure time is larger than  $t$  and  $RUL_i(t)$  is their remaining useful life starting from  $t$ . At the following time steps, the algorithm continues to update the estimate of  $\hat{RUL}(t) = MTTF(t)$ , until a fault is detected upon a deviation of the signal from its nominal value, outside its range of allowed variability; at this time, the RUL estimation algorithm is put to work to match the similarity



of the developing signal trajectory evolution to those in the reference library, and combine their failure times to provide an estimate of the RUL.

From the above explanation, it is clear that fault detection plays a relevant role in the RUL estimation procedure and structures it so that the RUL estimations are made up of two parts: MTTF-based part, first and RUL matching-based part, afterwards. The dividing instance is the fault detection: when a fault is detected, from then on the test signal trajectory is compared with those of reference to estimate the RUL; before the detection, the RUL is estimated as the MTTF. In spite of the importance of fault detection, in this work this is not treated specifically and the focus is kept on the RUL estimation task after detection. Future works will address the task of fault detection within the proposed approach and with respect to two particular developments: i) the possibility of skipping the fault detection process and proceeding from the beginning with the matching of the evolving test trajectories with the reference trajectories stored in the database: it is foreseeable that before the fault, the test trajectory will match well with the non-faulty segments of the reference trajectories and the estimated RUL will approximate the MTTF; thus, apart from increased computational and storage memory burdens the results are expected to not degrade significantly with respect to those obtained in this work; ii) the possibility of including in the computation of the RUL the accuracy of the fault detection method in identifying a deviation of the test trajectory from the nominal pattern evolution at the early stage as well as account for the reliability of fault detection.

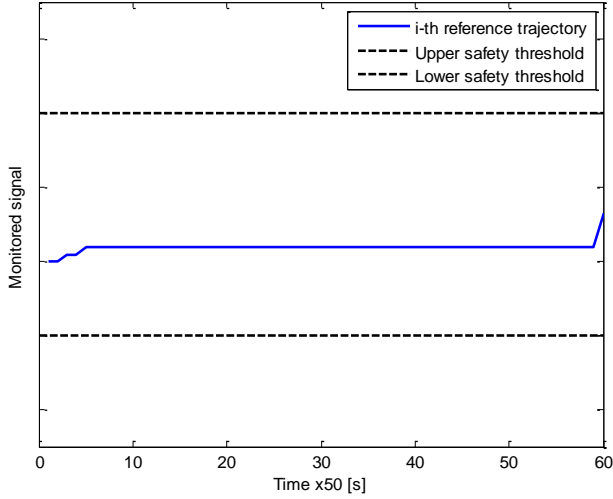


Figure 2 A reference signal trajectory on a time horizon  $T=3000$  [s]

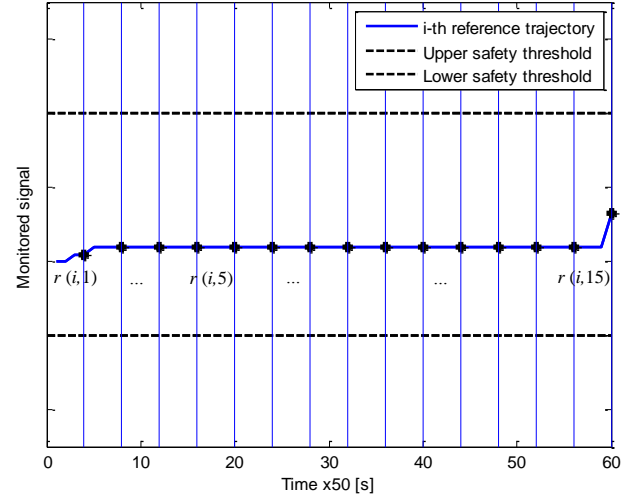


Figure 3 Reference signal trajectory of Figure 2 partitioned into  $k=15$  segments of length  $n=3000/15=200$  [s],  $j=1,2,...,15$

- **Step 2: trajectory pointwise difference computation.** At the current time  $t$ , the latest  $n$ -long segment of values of the test trajectory pattern  $\tilde{f}(t) = f(t-n+1, t-n+2, \dots, t)$  is normalized in  $[0.2, 0.8]$ . The pointwise difference  $\delta(\bullet)$  between the  $n$  values of pattern  $\tilde{f}(t)$  and those of the reference trajectory segment  $r(i, j)$  is computed:

$$\delta(i, j) = \tilde{f}(t) - r(i, j), \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, k \quad (2)$$

The matrix  $\bar{\delta}_{[N \times k]}$  contains the difference measures  $\delta(i, j)$  between all  $n$ -long segments of the reference trajectories and the test trajectory pattern of the monitored signal.

- **Step 3: trajectory pointwise similarity and distance score computation.** Classically, similarity is decided crisply: depending on whether the distance or similarity measure between two objects exceeds a specified threshold, the objects are classified into distinct categories of ‘similar’ or ‘non-similar’ [Angstenberger, 2001]. Such binary classification is restrictive when the situation is not so clear-cut and imprecision and uncertainty in similarity classification exist [Zimmermann et al., 1985]. In practice, there are numerous cases in which the similarity measure should allow for a gradual transition between ‘similar’ and

‘non-similar’ [Binaghi et al., 1993; Joentgen et al., 1999]. This can be achieved by resorting to a fuzzy logic modeling paradigm in which the pointwise difference of two trajectories is judged for similarity with respect to an “approximately zero” fuzzy set (FS) specified by a function which maps the elements  $\delta(i, j)$  of the difference matrix  $\overline{\overline{\delta}}_{[N \times k]}$  into their values  $\mu(i, j)$  of membership to the condition of “approximately zero”. The distance score  $d(i, j)$  between two trajectory segments is then computed as:

$$d(i, j) = 1 - \mu(i, j), \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, k \quad (3)$$

Common membership functions can be used for the definition of the FS, e.g. triangular, trapezoidal, and bell-shaped [Dubois et al., 1988]. In the exploration of these membership functions defined on a fixed support of the fuzzy set “approximately zero” and centered in the value zero, no particular sensitivity of the RUL estimation results has been recorded with respect to the particular shape of the function; in the application illustrated in the work, the following bell-shaped function has turned out to give more robust results due to its gradual smoothness:

$$\mu(i, j) = e^{-\left(\frac{-\ln(\alpha)}{\beta^2} \delta^2(i, j)\right)} \quad (4)$$

The arbitrary parameters  $\alpha$  and  $\beta$  can be set by the analyst to shape the desired interpretation of similarity into the fuzzy set: the larger the value of the ratio  $\frac{-\ln(\alpha)}{\beta^2}$ , the narrower the fuzzy set and the stronger the definition of similarity (Figure 4). The choice of the values of  $\alpha$  and  $\beta$  depends on the application; one may proceed to determining the value  $\beta$  of the difference value  $\delta$  which must have a degree of membership  $\mu$  equal to  $\alpha$  [Angstenberger, 2001]

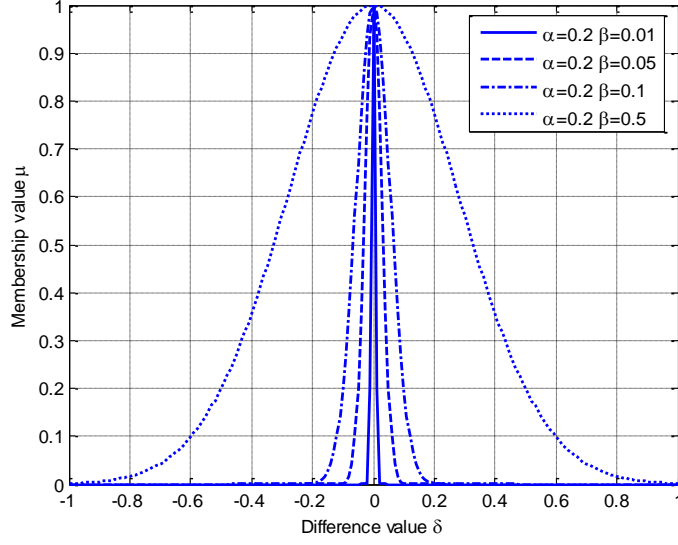


Figure 4 Bell-shaped membership functions with different values of parameter  $\beta$

- **Step 4: weight definition.** The  $\hat{RUL}(t)$  is estimated as a similarity-weighted sum of the  $RUL_i(t)$ :

$$\hat{RUL}(t) = \sum_{i|t_{fi} > t} w_i \cdot RUL_i(t), \quad i = 1, 2, \dots, N \quad (5)$$

The ideas behind the weighting of the individual  $RUL_i(t)$  is that: *i*) all failure trajectories in the reference library bring useful information for determining the RUL of the trajectory currently developing; *ii*) those segments of the reference trajectories which are most similar to the most recent segment of length  $n$  of the currently developing failure trajectory should be more informative in the extrapolation of the occurring trajectory to failure.

To assign the weight  $w_i$ , the minimum distance  $d_i^*$  along the  $i^{th}$  row of the matrix of Eq. (3) is first identified:

$$d_i^* = \min_{j=1, \dots, k} d(i, j), \quad i = 1, 2, \dots, N \quad (6)$$

The weight  $w_i$  is then computed, resorting to the arbitrarily chosen decreasing monotone function:

$$w_i = (1 - d_i^*) \cdot e^{\left(-\frac{1}{\beta} d_i^*\right)}, \quad i = 1, 2, \dots, N \quad (7)$$

The value of  $\beta$  in (7) is arbitrarily taken equal to that used in Eq. (4), for reducing the burden of the parametric sensitivity analysis performed for identifying the parameters most affecting the results.

Then, the weight  $w_i$  is normalized by:

$$w_i = w_i / \sum_{e=1}^N w_e \quad (8)$$

As shown in Figure 4, the smaller the minimum distance the larger the weight given to the  $i$ -th trajectory.

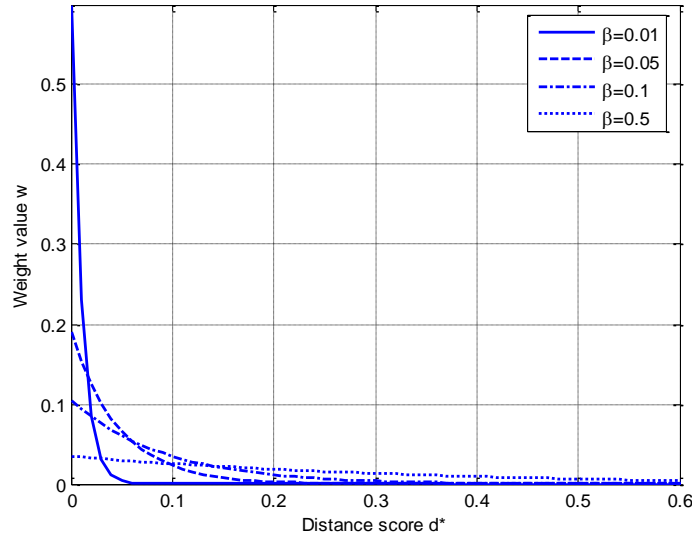
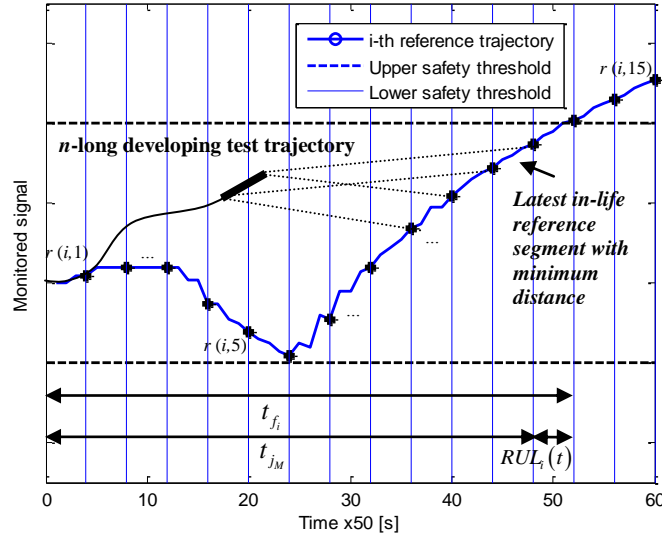


Figure 5 Weight vs. distance: the smaller the distance, the larger the weight value

- **Step 5:  $RUL_i(t)$  and  $RUL(t)$  estimation.** With respect to the generic  $i$ -th trajectory in the library for which  $t_{f_i} > t$ , the value  $RUL_i(t)$  is determined as:

$$RUL_i(t) = t_{f_i} - t_{j_M}, \quad i = 1, 2, \dots, N \quad (9)$$

where  $t_{j_M} = n \cdot \max_j \left( \arg \left( \delta(i, j) = d_i^* \right) \right)$  is the final time index of the latest-in-life segment of the  $i$ -th trajectory among those which has minimum distance  $d_i^*$  from the developing test trajectory ( $n$  is the test trajectory pattern length,  $\max_j \left( \arg \left( \delta(i, j) = d_i^* \right) \right)$  gives the largest column index  $j$  of  $r(i, \bullet)$  whose element is equal to  $d_i^*$ ). Thus,  $RUL_i(t)$  is the remaining time before failure on the reference trajectory starting from the end time of the latest-in-life segment of minimum distance from the developing trajectory (Figure 6).



**Figure 6** The remaining time before failure  $RUL_i(t)$  on the reference trajectory of Figure 2 starting from the end time of the latest-in-life segment of minimum distance from the occurring trajectory

This allows a conservative RUL estimation, biased towards “pessimistic” predictions of the RUL, because in the case that more than one segment along the  $i$ -th reference trajectory is closest to the developing test trajectory, the latest one is taken, i.e. the one closest to failure. Then, the estimate  $\hat{RUL}(t)$  of the remaining useful life along the developing trajectory is simply computed as in Eq. (5) with weights  $w_i$  evaluated by Eq. (8).

### 3. The LBE-XADS

The Lead-Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) is a sub-critical, fast reactor in which the fission process for providing thermal power  $P(t)$  is sustained by an external neutron source through spallation reaction by a proton beam  $Q(t)$  accelerated by a synchrotron on a lead-bismuth eutectic target [Bowman et al., 1992; Carminati et al., 1993; Rubbia et al., 1995; Van Tuyle et al., 1993; Venneri et al., 1993]. A simplified scheme of the plant is sketched in Figure 7. The primary cooling system is of pool-type with Lead-Bismuth Eutectic (LBE) liquid metal coolant leaving the top of the core, at full power nominal conditions, at temperature  $\tau_{LB}^{C,P}$  equal to 400 °C and then re-entering the core from the bottom through the down-comer at temperature  $\tau_{LB}^{P,C}$  equal to 300 °C. The average in-core temperature of the LBE  $T_{LB}^{av,C}$  is taken as the mean of  $\tau_{LB}^{C,P}$  and  $\tau_{LB}^{P,C}$ . The secondary cooling system is a flow of an organic diathermic oil at 290-320 °C, at full power conditions. Cooling of the diathermic oil in each loop is obtained through an air flow  $\Gamma_a(t)$  provided by three air coolers connected in series.

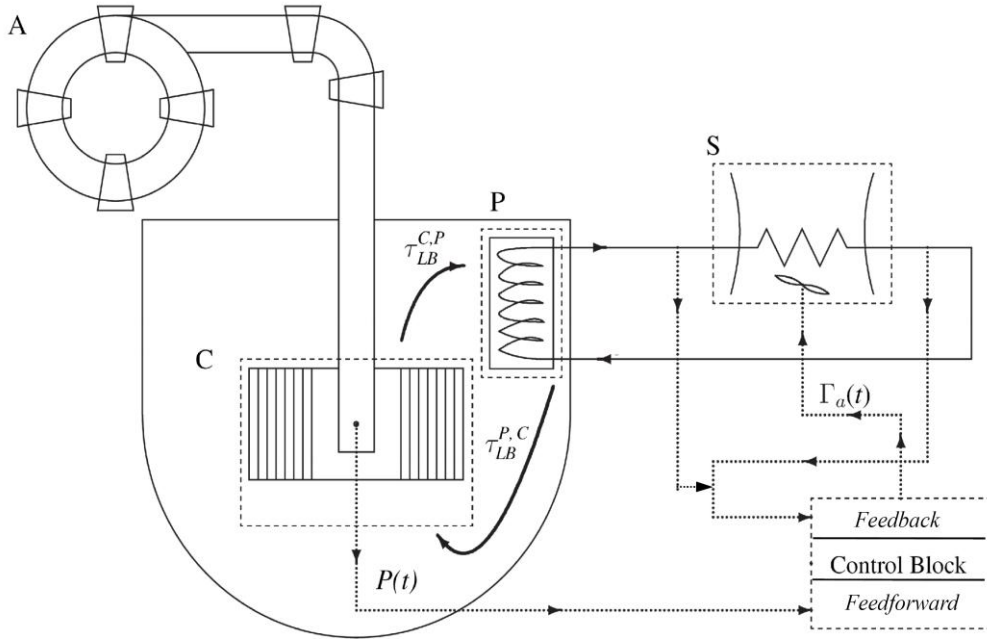


Figure 7 LBE-XADS simplified schematics. A = Accelerator; C = core; P = primary heat exchanger; S = secondary heat exchanger

A dedicated, dynamic simulation model has been implemented in SIMULINK for providing a simplified, lumped and zero-dimensional description of the coupled neutronic and thermo-hydraulic evolution of the system [Cammi et al., 2006]. The model allows the simulation of the system controlled dynamics as well as of the free dynamics when the control module is deactivated and the air cooler flow is kept constant.

Both feedforward and feedback digital control schemes have been adopted for the operation of the system. The control is set to keep a steady state value of approximately 300 °C of the average temperature of the diathermic oil  $T_o^{av,S}$ : this value represents the optimal working point of the diathermic oil at the steady state, full nominal power of 80 MWth. On the contrary, an oil temperature beyond the upper threshold  $T_o^{th,u}=340$  °C would lead to degradation of its physical and chemical properties, whereas a temperature below the lower threshold  $T_o^{th,l}=280$  °C could result in thermal shocks for the primary fluid and, eventually, for the structural components [Cammi et al., 2006].

Multiple component failures can occur during the system life. To simulate this, the model has been embedded within a Monte Carlo (MC) sampling procedure for injecting faults at random times and of random magnitudes. Samples of component failures are drawn within a mission time of 3000 [s]. The set of faults considered are:

- The PID controller fails stuck, with a random flow rate output value  $m_1$  sampled from a uniform distribution in [0,797] [kg/s].
- The air coolers fail stuck in a random position that provides a corresponding air flow mass  $m_2$  uniformly distributed in [0,1000] [kg/s].
- The feedforward controller fails stuck with a corresponding flow rate value  $m_3$  uniformly distributed in [0,797] [kg/s].

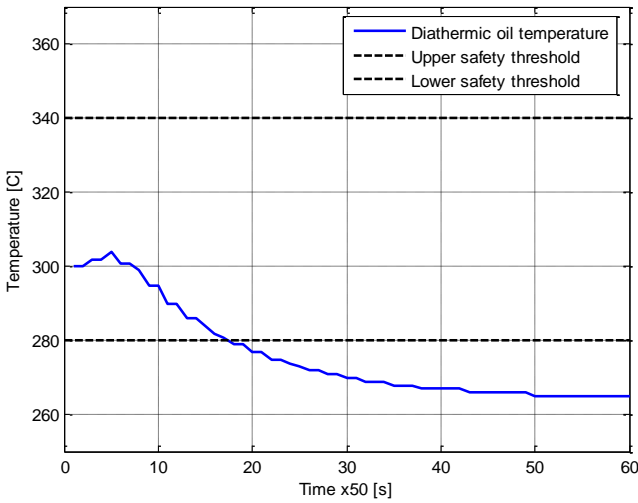


- The communication between air coolers actuators and PID controller fails so that the PID is provided with the same input value of the previous time step.

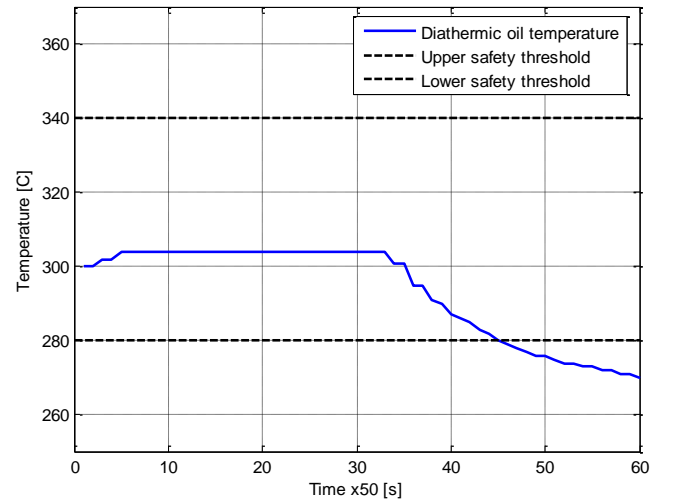
The sequence of multiple failures is generated by sampling the first failure time from the uniform distribution  $[0, 3000]$  [s] and the successive failure times from the conditional distributions, uniform from the last sampled time to 3000 [s]. This assumption is conservative, favoring larger number of failures in the sequence.

The evolution of the accident scenarios may lead to three different end states, within the mission time of 3000 [s] (Figures 8-12):

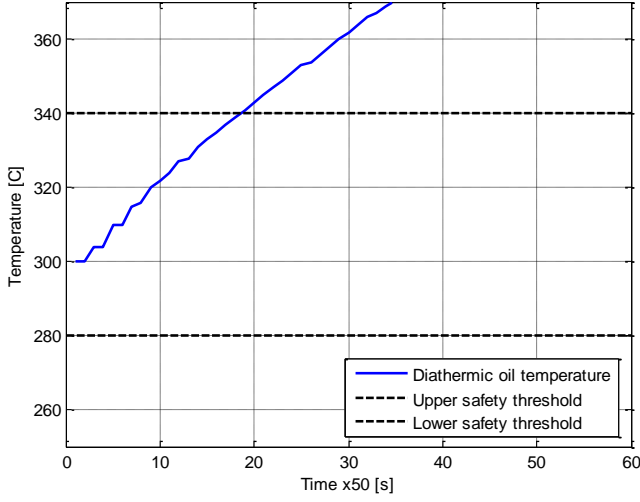
1. Low-temperature failure mode ( $T_o^{av,S} < T_o^{th,l}$ )
2. Safe mode ( $T_o^{th,l} < T_o^{av,S} < T_o^{th,u}$ )
3. High-temperature failure mode ( $T_o^{av,S} > T_o^{th,u}$ )



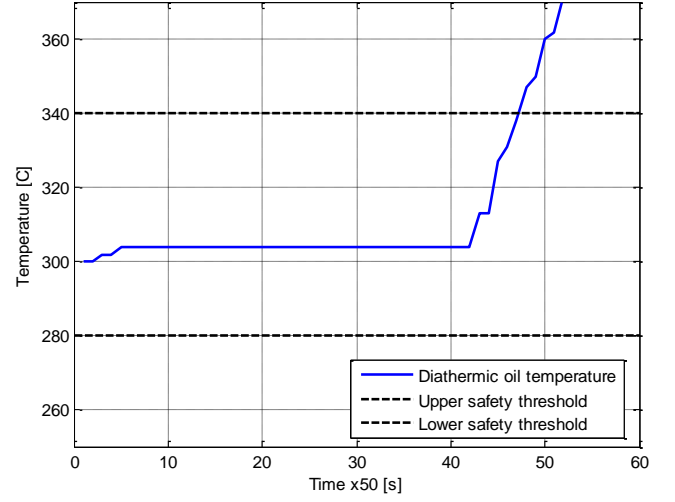
**Figure 8** Transient trajectory belonging to the low-temperature failure mode. System fails shortly



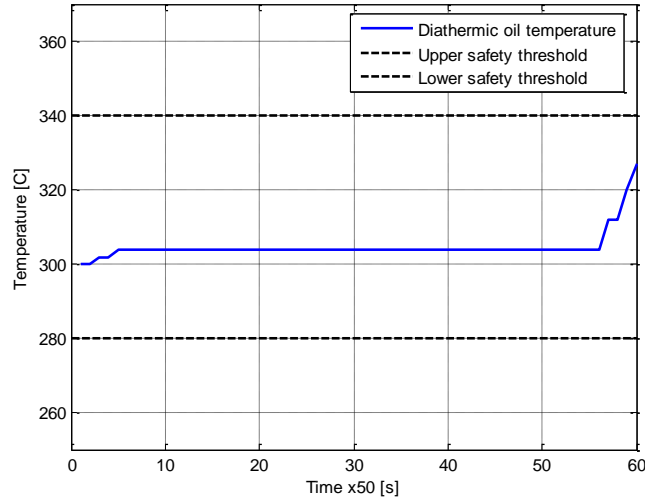
**Figure 9** Transient trajectory belonging to the low-temperature failure mode. System fails at later time



**Figure 10** Transient trajectory belonging to the high-temperature failure mode. System fails shortly



**Figure 11** Transient trajectory belonging to the high-temperature failure mode. System fails at later time



**Figure 12** Transient in which the system does not fail within the time horizon considered, although an accident has occurred

## 4. Results

### 4.1 Application of the $RUL(t)$ estimation procedure

The application of the procedure for computing the estimate  $\hat{RUL}(t)$  (Section 2) is hereafter illustrated; the goal is to demonstrate its feasibility for fast and reliable data-driven prognostics which can serve as operator support in emergency accident management. A database of reference trajectories for  $N = 6400$  failure scenarios (differing in faulty components, time of faults occurrence

and faults magnitude) is organized in the reference matrix  $\overline{R}_{[N \times k]}$ , where  $N = 6400$  and  $k = \frac{T}{n} = 30$ , respectively. The elements  $r(i, j)$  of the reference matrix will be compared for similarity with a test trajectory pattern containing the values of the temperature signal of the latest 100 time steps of the trajectory. For each of the test trajectories the procedural steps are performed as follows:

***Step 1: fault detection.***

The signal  $f(t)$  monitored starting from time  $t = 1$  [s] to the mission time  $T = 3000$  [s], with time step of 1 [s], is the LBE primary coolant temperature. This means that the actual safety parameter driving the RUL, i.e., the diathermic oil temperature, is followed only indirectly through the LBE temperature values, thus challenging further the robustness of the approach with respect to non-observable safety parameters. Obviously, other choices for  $f(t)$  may also be valid to drive the estimation, e.g. the diathermic oil temperature directly, another indirect parameter like the air flow or any combination of the observable parameters. At each time step  $t$ , its value is appended and stored in the vector containing the  $n - 1 = 99$  values of the signal collected at the previous times. The Mean Time to Failure  $MTTF(t)$  is calculated resorting to Eq. (1) and  $\hat{RUL}(t)$  is set equal to  $MTTF(t)$  for each time step, until any component failure is detected; the fault detection activates the on-line fuzzy based data-driven RUL estimation.

***Step 2: trajectory pointwise difference computation.***

The pointwise difference  $\delta(i, j)$  between the test pattern  $\tilde{f}(t) = f(t - 99, t - 98, \dots, t)$  and  $r(i, j)$  is evaluated.

***Step 3: trajectory pointwise similarity and distance score computation.***

The pointwise differences  $\delta(i, j)$  are mapped into values of membership  $\mu(i, j)$  of the “approximately zero” FS. The bell-shaped function of Eq. (4) is taken with parameters values  $\alpha = 0.2$  and  $\beta = 0.01$ , implying strong sharpness in the FS and thus in the similarity requirement

(Figure 4). The distance scores  $d(i, j)$  are then computed by Eq. (3),  $i = 1, 2, \dots, 6400$ ,  $j = 1, 2, \dots, 30$ .

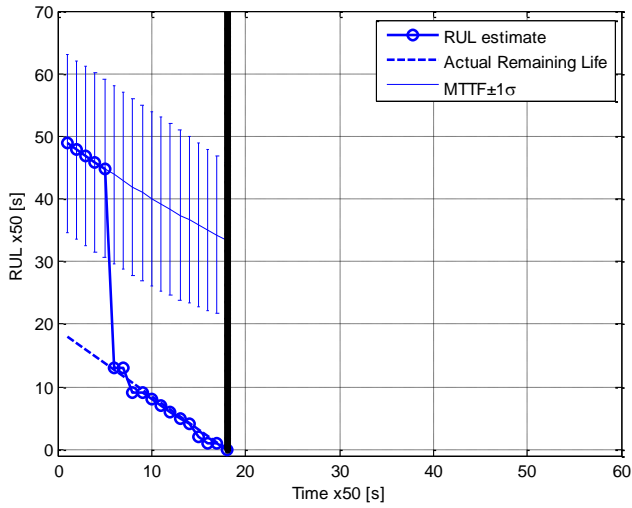
**Step 4: weight definition.**

The minimum distances  $d_i^*$  are evaluated (Eq. (6)), and the relative normalized weights  $w_i$  calculated through Eqs. (7) and (8),  $i = 1, 2, \dots, 6400$ .

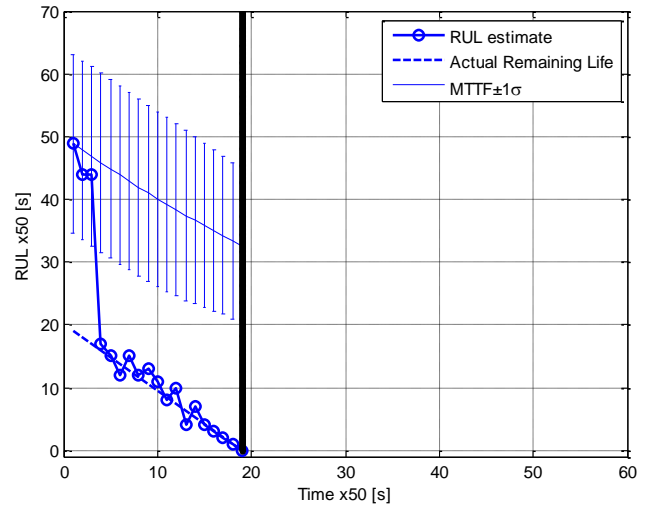
**Step 5:  $RUL_i(t)$  and  $RUL(t)$  estimation.**

For each reference trajectory in the library, an estimate  $\hat{RUL}_i(t)$  for the test trajectory is computed (Eq. (9),  $i = 1, 2, \dots, 6400$ ); then, the  $RUL_i(t)$  are aggregated in the weighted sum (Eq. (5)) with the weights  $w_i$  previously calculated.

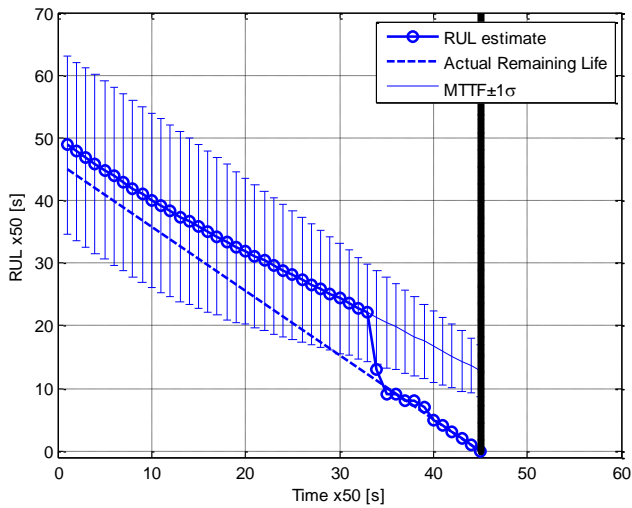
For the five trajectories sketched in Figures 8-12, the estimates of the  $MTTF(t)$  are plotted in Figures 13-17, in thin continuous lines with the bars of one standard deviation of the samples  $(t_{f_i} - t | t_{f_i} > t)$ , where  $t_{f_i}$  is the time at which the diathermic oil temperature profile of the  $i$ -th reference trajectory exceeds either thresholds  $T_o^{th,u}$  or  $T_o^{th,l}$ , with corresponding system loss of functionality. The  $\hat{RUL}(t)$  estimates obtained based on trajectory segments of  $n = 100$  [s] are plotted in bold circles; at the beginning of the test trajectories, the predictions match the  $MTTF(t)$ ; then, once a component failure is detected, the  $\hat{RUL}(t)$  estimate moves away from the  $MTTF(t)$  values towards the real  $RUL(t)$  (dashed thick line). In the Figures, the bold vertical line indicates the time of diathermic oil threshold exceedance. Notice that none of the estimates exceeds the actual failure time.



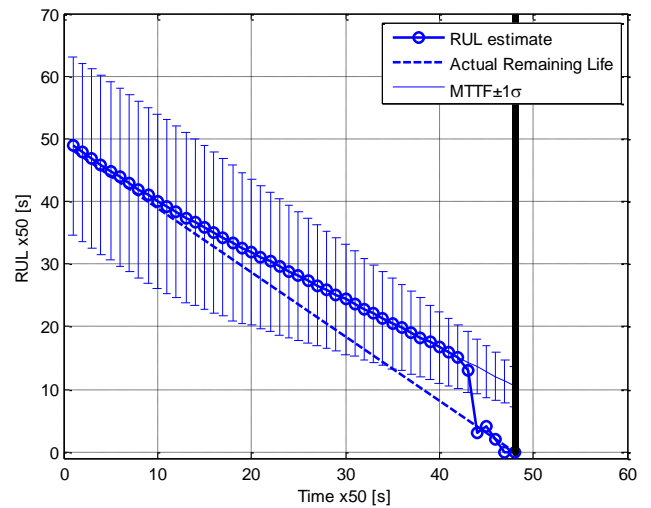
**Figure 13** RUL estimation for the trajectory of Figure 8 belonging to the low-temperature failure mode



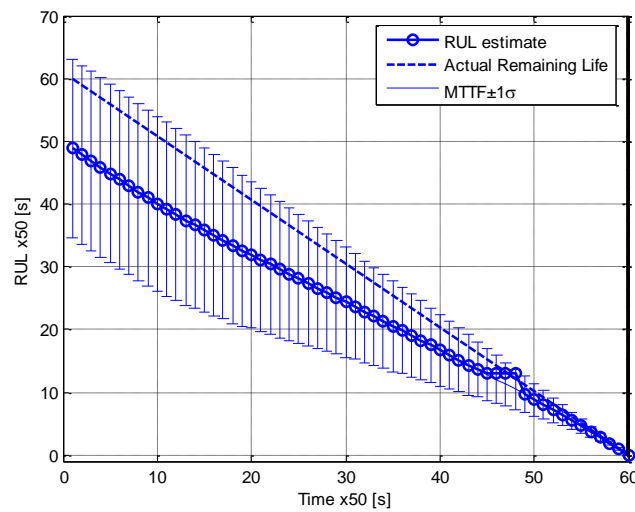
**Figure 15** RUL estimation for the trajectory of Figure 9 belonging to the high-temperature failure mode



**Figure 14** RUL estimation for the trajectory of Figure 10 belonging to the low-temperature failure mode



**Figure 16** RUL estimation for the trajectory of Figure 11 belonging to the high-temperature failure mode



**Figure 17** RUL estimation for the trajectory of Figure 12 which does not exceed any safety threshold value, although a failure sequence has occurred

## 4.2 Performance evaluation of the $RUL(t)$ estimation procedure

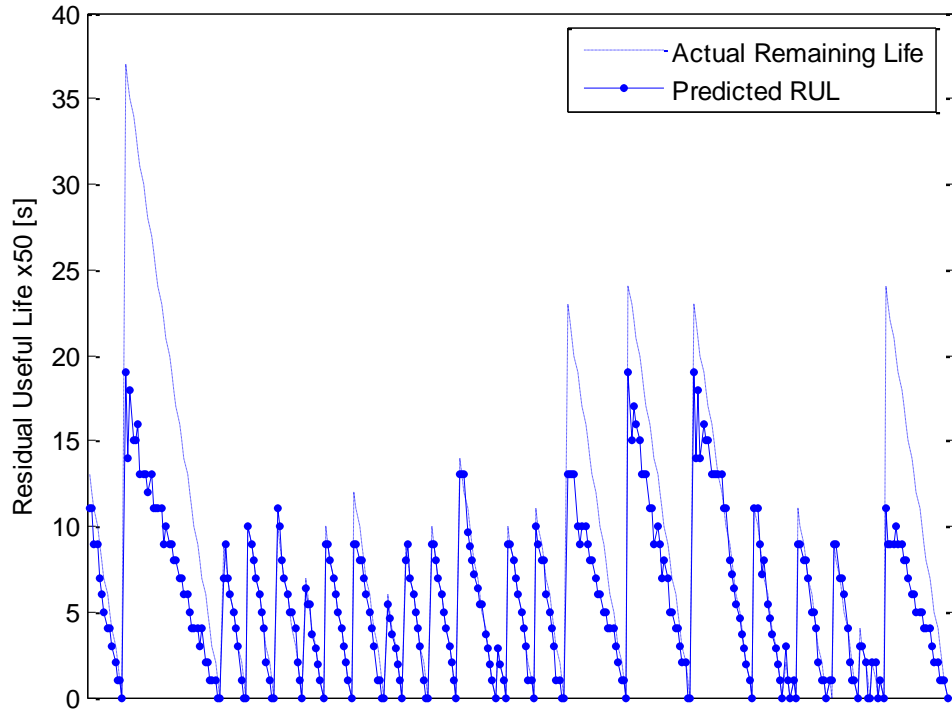
The performance of the RUL estimation procedure has been tested extensively on a batch of  $P=1280$  test trajectories, different from the reference trajectories. Figure 18 shows the  $\hat{RUL}(t)$  predictions (continuous line with dark bullets) compared to the actual remaining lives for 25 failure trajectories (graphically appended in sequence), after the fault has been detected (i.e., starting from step 2 of the algorithm). After fault detection, the initial predictions for each test trajectory tend to be similar, regardless of the eventual length of the test trajectory life duration because the initial deviation from the nominal evolution is little sensitive and only slightly moves the  $RUL(t)$  prediction away from the  $MTTF(t)$  value; this results in a conservative trend of initial anticipation of the RUL associated to trajectories whose failure actually occurs late in life. The largest RUL estimation errors occur for those trajectories in which the component failure is of low-magnitude, whose effect only slowly drives the system to failure and the prediction away from  $MTTF(t)$  towards the true  $RUL(t)$ .

To globally quantify the performance of the procedure, the mean relative error (RE) at time  $t$ , between the estimate  $\hat{RUL}(t)$  and its true value  $RUL(t)$ , is introduced:

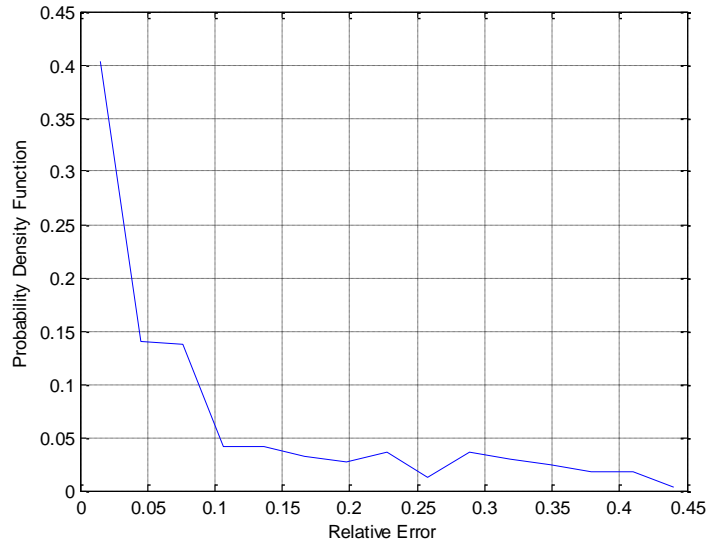
$$RE(t) = \frac{1}{P} \sum_{p=1}^P \frac{|\hat{RUL}_p(t) - RUL_p(t)|}{RUL_p(t)} \quad (10)$$

where  $RUL_p(t)$  is the actual remaining useful life at time  $t$  of test pattern  $p$ , and  $\hat{RUL}_p(t)$  its estimate,  $p = 1, 2, \dots, P$ .

Figure 19 shows the empirical probability density function of the mean relative error evaluated along 3000 [s] of evolution of the 1280 test trajectories. The distribution is highly skewed towards small values, with mean and median values equal to 0.09 and 0.05, respectively. This proves that the procedure most frequently makes small relative estimation errors.



**Figure 18 Predicted and actual remaining life for 25 test accident scenarios**

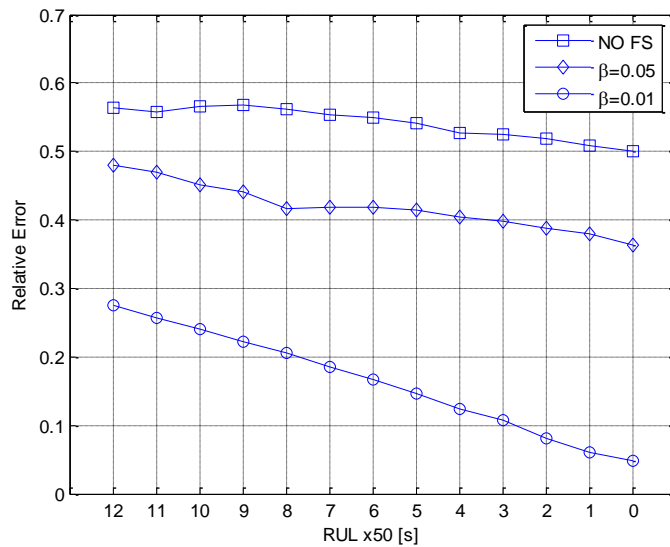


**Figure 19 Empirical probability density function of the relative errors on the 3000 [s] of the 1280 test trajectories**

Figure 20 shows the RE evaluated at time steps of 50 [s] in the last 600 [s] for different values of the parameter  $\beta$  in Eqs. (4) and (7), and including the case in which the FS “approximately zero” is not introduced in the step 3 of the procedure (i.e.,  $\beta=\infty$ ). It is seen that:

- the accuracy in the estimation of the RUL improves over time: as the RUL decreases, the performance index approximate its mean value 0.05.
- the performance degrades as the parameter  $\beta$  influencing the width of the fuzzy set “approximately zero” increases: the crisper the similarity definition, the worse the estimation performance.

For the application of the present work, the value of  $\beta = 0.01$  has turned out to provide satisfactory estimation results for all fault types, in few seconds on Intel® Centrino Core™2 Duo 1.80 MHz for one complete trajectory.



**Figure 20** Relative Error evaluated each 50 [s] starting from 600 [s] before failure, for 1280 test trajectories, with different pointwise similarity fuzziness

## 5. Conclusions

A similarity-based prognostics procedure for estimating the Remaining Useful Life (RUL) of a system has been proposed. Data from different transient failure scenarios are used to create a library of reference patterns of evolution. For estimating the RUL of a test pattern, its evolution data are matched to the patterns in the library and their known residual life times are used for the estimation, based on a fuzzy pointwise similarity concept.



The RUL estimation procedure involves four main steps: (1) computing the pointwise difference between test and reference patterns; (2) evaluating their fuzzy pointwise similarity and distance score; (3) defining the weights of the individual RUL estimates provided by the reference patterns; and (4) aggregating these to evaluate the system remaining useful life.

A case study regarding the RUL estimation for a large number of fault scenarios of the Lead Bismuth Eutectic eXperimental Accelerator Driven System (LBE-XADS) has demonstrated the promising potential of the procedure. The estimates of the expected RUL are refined as the failure transient develops in time. The results have been very satisfactory from the point of view of both accuracy of the estimation and computing time. The procedure may be used to allow predicting the remaining useful life before the future failure event with sufficient accuracy and timing for proactive maintenance/safety procedures purposes.

## References

- [Angstenberger, 2001] Angstenberger, L., Dynamic Fuzzy Pattern Recognition, International Series in Intelligent Technologies, 17, Kluwer Academic Publishers, 2001.
- [Barlett et al., 1992] Barlett, E.B., Uhrig, R.E., Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network, Nuclear Technology, 97, 1992.
- [Binaghi et al., 1993] Binaghi, E., Della Ventura, A., Rampini, A., Schettini, R. (1993) Fuzzy Reasoning Approach to Similarity Evaluation in Image Analysis. International Journal of Intelligent Systems, Vol. 8, 1993, p. 749-769, 1993.
- [Bowman et al., 1992] Bowman, C.D., Arthur, E.D., Lisowski, P.W., Lawrence, G.P., Jensen, R.J., Anderson, J.L., Blind, B., Cappiello, M., Davidson, J.W., England, T.R., Engel, L.N., Haight, R.C., Hughes, H.G., Ireland, J.R., Krakowski, R.A., LaBauve, R.J., Letellier, B.C., Perry, R.T., Russell, G.J., Staudhammer, K.P., Versamis, G., Wilson, W.B., Nuclear energy generation and waste transmutation using an accelerator-driven intense thermal neutron source. Nucl. Instr. Meth. Phys. Res. A 320, 336, 1992.
- [Cammi et al., 2006] Cammi, A., Luzzi, L., Porta, A. A. and Ricotti, M. E., Modelling and control strategy of the Italian LBE-XADS, Progress in Nuclear Energy, Volume 48, Issue 6, Pages 578-589, 2006.

- [Campolucci et al., 1999] Campolucci, P., Uncini, A., Piazza, F. and Rao, B.D., On-Line Learning Algorithms of Locally Recurrent Neural Networks, *IEEE Trans. Neural Networks*, 10, pp. 253-271, 1999.
- [Carminati et al., 1993] Carminati, F., Klapisch, R., Revol, J.P., Roche, Ch., Rubio, J.A., Rubbia, C., An Energy Amplifier for Cleaner and Inexhaustible Nuclear Energy Production Driven by a Particle Beam Accelerator. CERN Report CERN-AT-93-47(ET), 1993.
- [Chiang et al., 2001] Chiang L. H., Russel, E., Braatz R., Fault detection and diagnosis in industrial systems, Springer-Verlag, London, 2001.
- [Dragomir et al., 2007] Dragomir, O.E., Gouriveau, R., Zerhouni, N., Dragomir, F., Framework for a Distributed and Hybrid Prognostic System, In 4th IFAC Conference on Management and Control of Production and Logistics, 2007.
- [Dubois et al., 1988] Dubois, D., Prade, H., Testemale, C., Weighted Fuzzy Pattern Matching. *Fuzzy Sets and Systems*, 28, 1988, p. 313-331, 1988.
- [Gertsbakh, 2000] Gertsbakh, I., Reliability Theory With Applications to Preventive Maintenance. Springer: Berlin, 2000.
- [Glasstone et al., 1998] Glasstone, S, Sesonske, A., Nuclear reactor engineering. New Delhi: CBS Publishers and Distributors; 1998.
- [Jardine et al., 2006] Jardine, A.K.S, Lin, D., Banjevic, D., A review on machinery diagnostics and prognostics implementing condition based maintenance, *Mech. Syst. and Sig. Process.*, Vol. 20, pp. 1483-1510, 2006.
- [Joentgen et al., 1999] Joentgen, A., Mikenina, L., Weber, R., Zimmermann, H.-J., Dynamic Fuzzy Data Analysis based on Similarity between Functions. *Fuzzy Sets and Systems*, 105 (1), 1999, p. 81-90, 1999.
- [Peel et al., 2008] Peel, L., Data Driven Prognostics using a Kalman Filter Ensemble of Neural Network Models, International Conference on Prognostics and Health Management, 2008.
- [Rubbia et al., 1995] Rubbia, C., Rubio, J.A., Buono, S., Carminati, F., Fitier, N., Galvez, J., Gels, C., Kadi, Y., Klapisch, R., Mandrillon, P., Revol, J.P., Roche, Ch., Conceptual Design of a Fast Neutron Operated High Power Energy Amplifier. CERN Report CERN-AT-95-44(ET), 1995.
- [Santosh et al., 2009] Santosh, T.V., Srivastava, A., Sanyasi Rao, V.V.S., Gosh, A. K., Kushwaha, H.S., Diagnostic System for Identification of Accident Scenarios in Nuclear Power Plants using Artificial Neural Networks, *Reliability Engineering and System Safety*, 94, 759-762, 2009.
- [Van Tuyle et al., 1993] Van Tuyle, G.J., Todosow, M., Geiger, M.J., Aronson, A.L., Takahashi, H., Accelerator-driven subcritical target concept for transmutation of nuclear wastes. *Nucl. Technol.* 101, 1, 1993.

- [Venneri et al., 1993] Venneri, F., Bowman, C.D., Jameson R., Accelerator-driven Transmutation of Waste (ATW) - A New Method for Reducing the Long-term Radioactivity of Commercial Nuclear Waste. Los Alamos Report LA-UR-93-752, 1993
- [Wang et al., 2008] Wang T., Yu, J., Siegel, D., Lee, J., A similarity based prognostic approach for Remaining Useful life estimation of Engineered Systems, International Conference on Prognostics and Health Management, 2008.
- [Wang et al., 2004] Wang, W.Q., Goldnaraghi, M.F., Ismail, F., Prognosis of Machine Health Condition using Neuro-Fuzzy Systems, Mechanical Systems and Signal Processing, 18, 813-831, 2004.
- [Yan et al., 2004] Yan, J., Koç, M., Lee, J., A Prognostic Algorithm for Machine performance Assessment and its Application, Production Planning and Control, 15:8, 796-801, 2004.
- [Yuan et al., 1997] Yuan, B., Klir, G., Data driven identification of key variables, In: Ruan, D. (Ed.), Intelligent Hybrid Systems Fuzzy Logic, Neural Network, and Genetic Algorithms. Kluwer Academic Publishers, pp.161–187, 1997.
- [Zimmermann et al., 1985] Zimmermann, H.-J., Zysno, P., Quantifying Vagueness in Decision Models. European Journal of Operational Research, Vol. 22, 1985, p. 148-158, 1985.
- [Zadeh, 1965] Zadeh, L., Fuzzy sets, Inform. Control 8, 338–353, 1965.
- [Zio et al., 2005] Zio, E., Baraldi, P., Identification of nuclear transients via optimized fuzzy clustering, Annals of Nuclear Energy, 32, 1068–1080, 2005.